On Interference Networks with Feedback and Delayed CSI

Ravi Tandon Soheil Mohajer H. Vincent Poor Shlomo Shamai (Shitz)

Department of EE
Princeton University
Princeton, N.J.
{*rtandon,smohajer,poor*}*@princeton.edu*

Department of EE
Technion-Israel Inst. of Technology
Technion City, Haifa 32000, Israel
sshlomo@ee.technion.ac.il *

September 27, 2011

Abstract

The degrees of freedom (**DoF**) region of the two-user MIMO interference channel (IC) is completely characterized in the presence of noiseless channel output feedback from each receiver to its respective transmitter and with the assumption of delayed channel state information (CSI) at the transmitters. It is shown that having output feedback and delayed CSI at the transmitters can strictly enlarge the **DoF** region when compared to the case in which only delayed CSI is available at the transmitters. The proposed coding schemes that achieve the corresponding **DoF** region with feedback and delayed CSI utilize both resources, i.e., feedback and delayed CSI in a non-trivial manner. Furthermore, cases are identified in which output feedback and delayed CSI alone are sufficient to achieve the **DoF** region achievable with perfect, instantaneous CSI.

Also, the **DoF** region of the scalar $2 \times 2 \times 2$ two-hop interference network is characterized. It is shown that the total **DoF** for this network collapses from 2 (with instantaneous CSI) to 4/3 (with one-hop output feedback and delayed CSI).

1 Introduction

In many wireless networks, multiple pairs of transmitters/receivers wish to communicate over a shared medium. In such situations, due to the broadcast and superposition nature of the

*The work of R. Tandon, S. Mohajer and H. V. Poor was supported in part by the Air Force Office of Scientific Research under MURI Grant FA9550-09-1-0643, and in part by the DTRA under Grant HDTRA-07-1-0037. The work of S. Shamai was supported by the Israel Science Foundation (ISF), and the Philipson Fund for Electrical Power.

wireless medium, the effect of interference is inevitable. Hence, management of interference is of extreme importance in such networks.

Various interference management techniques have been proposed over the past few decades. The more traditional approaches to deal with interference either treat it as noise (in the low interference regime) or decode and then remove it from the received signal (in the high interference regime). However, such techniques are not strong enough to achieve the optimal performance of the network even in the simple interference channel with two pairs of multiple-input multiple-output (MIMO) transceivers. Recently, more sophisticated schemes, such as interference alignment [1,2] and (aligned) interference neutralization [3–5] have been proposed for managing interference, which can significantly increase the achievable rate over the interference networks. However, these techniques are usually based on availability of instantaneous (perfect) channel state information (p-CSI) at the transmitters. Such an assumption is perhaps not very realistic in practical systems, at least when dealing with fast fading links.

Quite surprisingly, it is shown by Maddah-Ali and Tse that even delayed (stale) CSI is helpful to improve the achievable rate of wireless network with multiple flows, even if the channel realizations are independent over time. In a significant paper [6], they showed that the sum Degrees of Freedom (**DoF**) of 4/3 is achievable for a broadcast MIMO network with two transmit antennas and one antenna at each receiver. This is in contrast to $\mathbf{DoF} = 1$, which is known to be optimal when no CSI is available. This technique is further studied in [7, 8], where the authors showed that the delayed CSI can also improve the achievable **DoF** of the X-channel.

The **DoF** region of the two-user MIMO interference channel with delayed CSI is completely characterized by Vaze and Varanasi [9,10]. They have shown that, depending on the parameters of the channel (the number of antennas at each terminal), the **DoF** with delayed CSI can be strictly better than that of no CSI, and worse than that with instantaneous CSI. The role of output feedback in the performance of wireless communication systems has received considerable attention over the past few decades. It is well known that feedback does not increase the capacity of point-to-point discrete memoryless channels. Unlike the point-to-point case, feedback can increase the capacity of the multiple-access and broadcast channels. The effects of feedback on the capacity region of the interference channel have been studied in several recent papers (see [11] and references therein).

One question that can be raised here is whether output feedback can be helpful with delayed CSI or not. For the case of the broadcast channel (BC), this question is answered in a negative way in [12]: having output feedback beside delayed CSI does not increase the **DoF** region of the MIMO BC.

In this work, we study this question for the two-user interference channel (IC), where each transmitter is provided with the past state information of the channel, as well as the received signal (feedback) from its respective receiver. It turns out that, surprisingly,

existence of output feedback can increase the **DoF** region of the interference channel. In the presence of output feedback with delayed CSI, transmitter 1 besides being able to reconstruct the interference it caused at receiver 2, can also reconstruct a part of the signal intended to receiver 2. This is in contrast to the case of the MIMO BC, where all information symbols are created at one transmitter, and hence output feedback in addition to delayed CSI does not increase the **DoF** region of the MIMO BC.

We also study the scalar $2\times2\times2$ two-hop interference network with the model of feedback and delayed CSI. The $2\times2\times2$ two-hop network is a concatenation of two 2-user interference channels. This model is a particularly relevant model for multi-hop interference networks and has recently received attention. In particular, reference [4] has shown that with perfect CSI, the optimal total **DoF** for the $2\times2\times2$ network is 2. In this paper, we completely characterize the **DoF** region for the $2\times2\times2$ network with feedback and delayed CSI. It is shown that the total **DoF** for this network collapses from 2 (with instantaneous CSI) to 4/3 (with one-hop output feedback and delayed CSI). The optimal coding scheme exhibits an interesting feature: coding over the first hop is done to mimic the X channel. This enables the relays to create a virtual MISO broadcast channel in the second hop to the receivers.

2 MIMO IC with Feedback and Delayed CSI

We consider the (M_1, M_2, N_1, N_2)-MIMO-IC with fast fading under the assumptions of (A-I) noiseless causal channel output feedback from each receiver to its respective transmitter and (A-II) the availability of delayed CSI at the transmitters (see Figure 1). We denote the transmitters by $\mathbf{Tx}_1$ and $\mathbf{Tx}_2$ and the receivers by $\mathbf{Rx}_1$ and $\mathbf{Rx}_2$. The number of antennas at transmitter m is denoted as M_m, $m = 1, 2$; the number of antennas at receiver n is denoted by N_n, $n = 1, 2$. The channel outputs at the receivers are given as

$$
\begin{aligned}
Y_1(t) &= \mathbf{H}_{11}(t)X_1(t) + \mathbf{H}_{12}(t)X_2(t) + Z_1(t)\\
Y_2(t) &= \mathbf{H}_{21}(t)X_1(t) + \mathbf{H}_{22}(t)X_2(t) + Z_2(t),
\end{aligned}
$$

where $X_m(t)$ is the signal transmitted by mth transmitter $\mathbf{Tx}_m$; $\mathbf{H}_{nm}(t) \in \mathbb{C}^{N_n\times M_m}$ denotes the channel matrix between nth receiver and mth transmitter; and $Z_n(t) \sim \mathcal{CN}(0, I_{N_n})$, for $n = 1, 2$, is the additive noise at receiver n. The power constraints are $\mathbb{E}||X_m(t)||^2 \leq P$, for $\forall\ m, t$.

We denote $\mathbf{H}(t) = \{\mathbf{H}_{11}(t), \mathbf{H}_{12}(t), \mathbf{H}_{21}(t), \mathbf{H}_{22}(t)\}$ as the collection of all channel matrices at time t. Furthermore, $\mathbf{H}^{t-1} = \{\mathbf{H}(1), \mathbf{H}(2), \ldots, \mathbf{H}(t-1)\}$ denotes the set of all channel matrices up till time $(t-1)$. Similarly, we denote $Y_n^{t-1} = \{Y_n(1), \ldots, Y_n(t-1)\}$ as the set of all channel outputs at receiver n up till time $(t-1)$. A coding scheme with block length T for

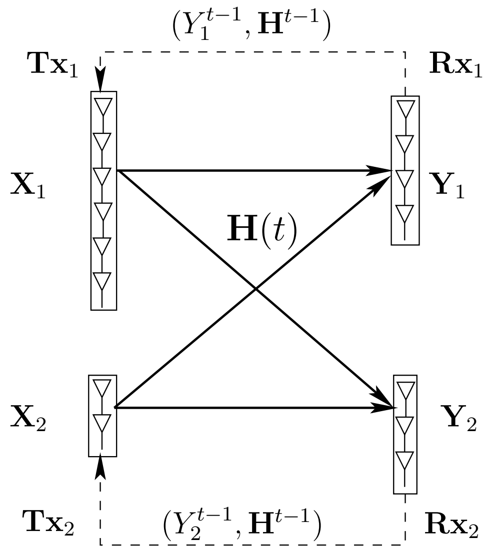

Figure 1: MIMO-IC with output feedback and delayed CSI.

the MIMO-IC with feedback and delayed CSI consists of a sequence of encoding functions:

$$X_1(t) = f_{1,t}^T \left(W_1, \mathbf{H}^{t-1}, Y_1^{t-1}\right)$$
$$X_2(t) = f_{2,t}^T \left(W_2, \mathbf{H}^{t-1}, Y_2^{t-1}\right),$$

defined for $t = 1, \ldots, T$, and two decoding functions:

$$\hat{W}_1 = g_1^T(Y_1^n, \mathbf{H}^n), \quad \hat{W}_2 = g_2^T(Y_2^n, \mathbf{H}^n).$$

A rate pair $(R_1(P), R_2(P))$ is achievable if there exists a sequence of coding schemes such that $\mathbb{P}(W_m \neq \hat{W}_m) \to 0$ as $T \to \infty$ for both $m = 1, 2$. The capacity region $\mathcal{C}(P)$ is defined as the set of all achievable rate pairs $(R_1(P), R_2(P))$ We define the DoF region as follows:

$$\begin{aligned} &\mathbf{DoF}^{\mathrm{FB,d-CSI}} \\ &= \Big\{ (d_1, d_2) \Big| d_m \geq 0, \text{ and } \exists (R_1(P), R_2(P)) \in \mathcal{C}(P) \\ &\qquad \text{s.t. } d_m = \lim_{P \to \infty} \frac{R_m(P)}{\log_2(P)}, m = 1, 2 \Big\}. \end{aligned} \tag{1}$$

We denote the **DoF** regions corresponding to the cases of no-CSI,, perfect-CSI, delayed-CSI, with output feedback, and with output feedback as well as delayed CSI as follows:

- No CSI: $\mathbf{DoF}^{\mathrm{No-CSI}}$

$$X_m(t) = f_{m,t}^T(W_m), m = 1, 2.$$

- Perfect CSI: $\mathbf{DoF}^{\text{p}-\text{CSI}}$

$$X_m(t) = f^T_{m,t}(W_m, \mathbf{H}^T), m = 1, 2.$$

- Delayed CSI: $\mathbf{DoF}^{\text{d}-\text{CSI}}$

$$X_m(t) = f^T_{m,t}(W_m, \mathbf{H}^{t-1}), m = 1, 2.$$

- Output Feedback: $\mathbf{DoF}^{\text{FB}}$

$$X_m(t) = f^T_{m,t}(W_m, Y_m^{t-1}), m = 1, 2.$$

- Output feedback and delayed CSI: $\mathbf{DoF}^{\text{FB},\text{d}-\text{CSI}}$

$$X_m(t) = f^T_{m,t}(W_m, Y_m^{t-1}, \mathbf{H}^{t-1}), m = 1, 2.$$

One of the contributions of this paper is a complete characterization of $\mathbf{DoF}^{\text{FB},\text{d}-\text{CSI}}$, stated in the following theorem:

Theorem 1 *The* **DoF** *region of the* (M_1, M_2, N_1, N_2)*-MIMO IC with channel output feedback and delayed CSI,* $\mathbf{DoF}^{\text{FB},\text{d}-\text{CSI}}$ *is given as the set of all non-negative pairs* (d_1, d_2) *that satisfy*

$$\begin{aligned}
d_1 &\leq \min(M_1, N_1) && (2)\\
d_2 &\leq \min(M_2, N_2) && (3)\\
d_1 + d_2 &\leq \min\Big\{M_1 + M_2, N_1 + N_2, \max(M_1, N_2), \max(M_2, N_1)\Big\} && (4)\\
\frac{d_1}{\min(N_1 + N_2, M_1)} + \frac{d_2}{\min(N_2, M_1)} &\leq \frac{\min(N_2, M_1 + M_2)}{\min(N_2, M_1)} && (5)\\
\frac{d_1}{\min(N_1, M_2)} + \frac{d_2}{\min(N_1 + N_2, M_2)} &\leq \frac{\min(N_1, M_1 + M_2)}{\min(N_1, M_2)}. && (6)
\end{aligned}$$

For comparison, we recall the DoF region with perfect, instantaneous CSI at the transmitters $\mathbf{DoF}^{\text{p}-\text{CSI}}$ [13]:

$$\begin{aligned}
d_1 &\leq \min(M_1, N_1) && (7)\\
d_2 &\leq \min(M_2, N_2) && (8)\\
d_1 + d_2 &\leq \min\Big\{M_1 + M_2, N_1 + N_2, \max(M_1, N_2), \max(M_2, N_1)\Big\}. && (9)
\end{aligned}$$

In addition, the **DoF** region with delayed CSI, $\mathbf{DoF}^{\text{d}-\text{CSI}}$ was characterized in [9]. This region is given by the set of inequalities as in Theorem 1 along with two more inequalities.

In particular, to characterize $\mathbf{DoF}^{\text{d}-\text{CSI}}$, [9] defines two mutually exclusive conditions:

$$\text{Condition 1}: M_1 > N_1 + N_2 - M_2 > N_1 > N_2 > M_2 > N_2\left(\frac{N_2 - M_2}{N_1 - M_2}\right) \tag{10}$$

$$\text{Condition 2}: M_2 > N_1 + N_2 - M_1 > N_2 > N_1 > M_1 > N_1\left(\frac{N_1 - M_1}{N_2 - M_1}\right). \tag{11}$$

If condition 1 holds, then $\mathbf{DoF}^{\text{d}-\text{CSI}}$ is given by the inequalities in Theorem 1 and the following additional bound (bound L_4 in [9]):

$$d_1 + d_2\left(\frac{N_1 + 2N_2 - M_2}{N_2}\right) \leq N_1 + N_2. \tag{12}$$

If condition 2 holds, then $\mathbf{DoF}^{\text{d}-\text{CSI}}$ is given by the inequalities in Theorem 1 and the following additional bound (bound L_5 in [9]):

$$d_2 + d_1\left(\frac{N_2 + 2N_1 - M_1}{N_1}\right) \leq N_1 + N_2. \tag{13}$$

Hence, from Theorem 1, we have the following relationship:

$$\mathbf{DoF}^{\text{No}-\text{CSI}} \subseteq \mathbf{DoF}^{\text{d}-\text{CSI}} \subseteq \mathbf{DoF}^{\text{FB},\text{d}-\text{CSI}} \subseteq \mathbf{DoF}^{\text{p}-\text{CSI}}.$$

Converse proof for Theorem 1: the upper bounds (2)-(3) are straightforward from the point-to-point MIMO channel. The sum degrees of freedom bound in (4) can be proved in a manner similar to the proof of [13]. Hence, to show that $\mathbf{DoF}^{\text{FB},\text{d}-\text{CSI}}$ is contained in the region given by (2)-(4), we need only to prove the bounds (5) and (6) for the case of output feedback and delayed CSI. Since (5) and (6) are symmetric, we need to prove that if $(d_1, d_2) \in \mathbf{DoF}^{\text{FB},\text{d}-\text{CSI}}$, then (d_1, d_2) must satisfy the bound (5). The proof of bound (5) closely follows the converse proof in [9]. We establish the bound (5) in Section 7.1.

Coding schemes with feedback and delayed CSI that achieve the region $\mathbf{DoF}^{\text{FB},\text{d}-\text{CSI}}$ stated in Theorem 1 are presented in Section 4.

3 $2 \times 2 \times 2$ Network with Feedback and Delayed CSI

We consider the $2 \times 2 \times 2$ interference network with fast fading under the assumptions of (A-I) noiseless causal channel output feedback from relay n to transmitter n, and receiver n to relay n, $n = 1, 2$ and (A-II) the availability of delayed (one-hop) CSI at the transmitters and at the relays. We denote the transmitters (source) by S_1 and S_2, relays by R_1 and R_2, and the receivers (destination) by D_1 and D_2 (see Figure 2).

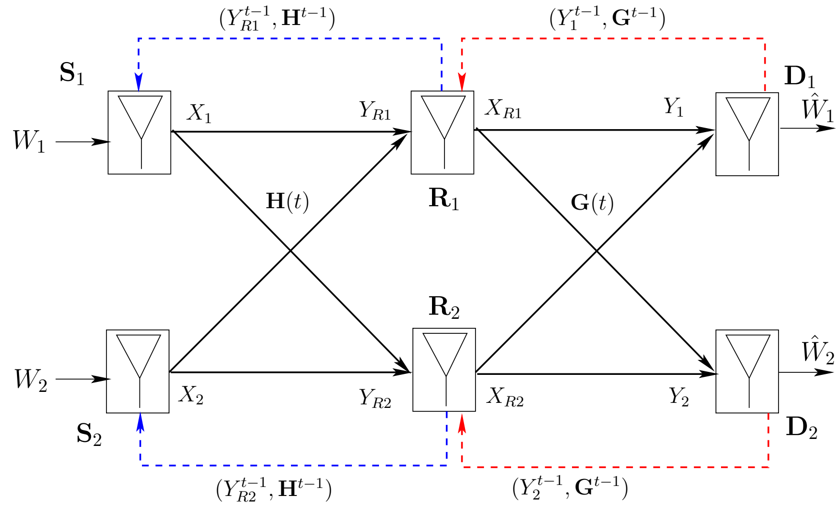

Figure 2: The $2 \times 2 \times 2$ network with feedback and delayed CSI.

The channel outputs at the relays are given as

$$Y_{R1}(t) = H_{11}(t)X_1(t) + H_{12}(t)X_2(t) + Z_1(t) \tag{14}$$

$$Y_{R2}(t) = H_{21}(t)X_1(t) + H_{22}(t)X_2(t) + Z_2(t), \tag{15}$$

and the channel outputs at the receivers are given as

$$Y_1(t) = G_{11}(t)X_{R1}(t) + G_{12}(t)X_{R2}(t) + U_1(t) \tag{16}$$

$$Y_2(t) = G_{21}(t)X_{R1}(t) + G_{22}(t)X_{R2}(t) + U_2(t), \tag{17}$$

where $X_m(t)$ is the signal transmitted by mth transmitter S_m; $X_{Rm}(t)$ is the signal transmitted by mth relay R_m; $H_{nm}(t) \in \mathbb{C}$ denotes the channel gain between nth relay and mth transmitter; $G_{nm}(t) \in \mathbb{C}$ denotes the channel gain between nth receiver and mth relay; and $Z_n(t), U_n(t) \sim \mathcal{CN}(0,1)$, for $n = 1, 2$, are the additive noises at relay n and receiver n, respectively. The power constraints are $\mathbb{E}||X_m(t)||^2 \leq P$, for $\forall\ m, t$.

We denote $\mathbf{H}(t) = \{H_{11}(t), H_{12}(t), H_{21}(t), H_{22}(t)\}$ as the collection of all channel coefficients of the first hop at time t; similarly, we denote $\mathbf{G}(t) = \{G_{11}(t), G_{12}(t), G_{21}(t), G_{22}(t)\}$ as the collection of all channel coefficients of the second hop at time t. Furthermore, $\mathbf{H}^{t-1} = \{\mathbf{H}(1), \mathbf{H}(2), \ldots, \mathbf{H}(t-1)\}$ denotes the set of all channel coefficients (of first hop) up till time $(t-1)$. Similarly, we denote $Y_n^{t-1} = \{Y_n(1), \ldots, Y_n(t-1)\}$ as the set of all channel outputs at receiver n up till time $(t-1)$. A coding scheme with block length T for the MIMO-IC with feedback and delayed CSI consists of a sequence of encoding functions at the transmitters:

$$X_1(t) = f_{1,t}^T \left(W_1, \mathbf{H}^{t-1}, Y_{R1}^{t-1}\right) \tag{18}$$

$$X_2(t) = f_{2,t}^T \left(W_2, \mathbf{H}^{t-1}, Y_{R2}^{t-1}\right), \tag{19}$$

a sequence of encoding functions at the relays:

$$X_{R1}(t) = g_{1,t}^{T}\left(\mathbf{G}^{t-1}, \mathbf{H}^{t-1}, Y_{R1}^{t-1}, Y_{1}^{t-1}\right) \tag{20}$$

$$X_{R2}(t) = g_{2,t}^{T}\left(\mathbf{G}^{t-1}, \mathbf{H}^{t-1}, Y_{R2}^{t-1}, Y_{2}^{t-1}\right), \tag{21}$$

defined for $t = 1, \dots, T$, and two decoding functions:

$$\hat{W}_1 = \psi_1^T(Y_1^n, H^n) \quad \hat{W}_2 = \psi_2^T(Y_2^n, H^n). \tag{22}$$

For this network, we denote the **DoF** region with perfect CSI by $\mathbf{DoF}_{222}^{\mathrm{p-CSI}}$ and with feedback and delayed CSI by $\mathbf{DoF}_{222}^{\mathrm{FB,d-CSI}}$. We state our second result in the following theorem.

Theorem 2 *The* **DoF** *region of the* $2 \times 2 \times 2$ *network with channel output feedback and delayed CSI,* $\mathbf{DoF}_{222}^{\mathrm{FB,d-CSI}}$ *is given as the set of all non-negative pairs* (d_1, d_2) *that satisfy*

$$\frac{d_1}{2} + d_2 \leq 1 \tag{23}$$

$$d_1 + \frac{d_2}{2} \leq 1. \tag{24}$$

For comparison, we recall the DoF region with perfect, instantaneous CSI at the transmitters $\mathbf{DoF}_{222}^{\mathrm{p-CSI}}$ [4]:

$$d_1 \leq 1 \tag{25}$$

$$d_2 \leq 1. \tag{26}$$

Figure 3 shows these two regions.

The converse proof for Theorem 2 follows from the following set of arguments. Consider an enhanced system in which both messages (W_1, W_2) are available at both the relays. In this enhanced system, we have a $(2, 1, 1)$-MISO broadcast channel from the (co-located) relays to the two receivers. We know the **DoF** region for this enhanced MISO broadcast channel with output feedback and delayed CSI which is given as [12]:

$$\frac{d_1}{2} + d_2 \leq 1 \tag{27}$$

$$d_1 + \frac{d_2}{2} \leq 1. \tag{28}$$

We complete the proof of Theorem 2 in Section 5 by showing the achievability of the point $(2/3, 2/3)$ for the $2 \times 2 \times 2$ network with feedback and delayed CSI.

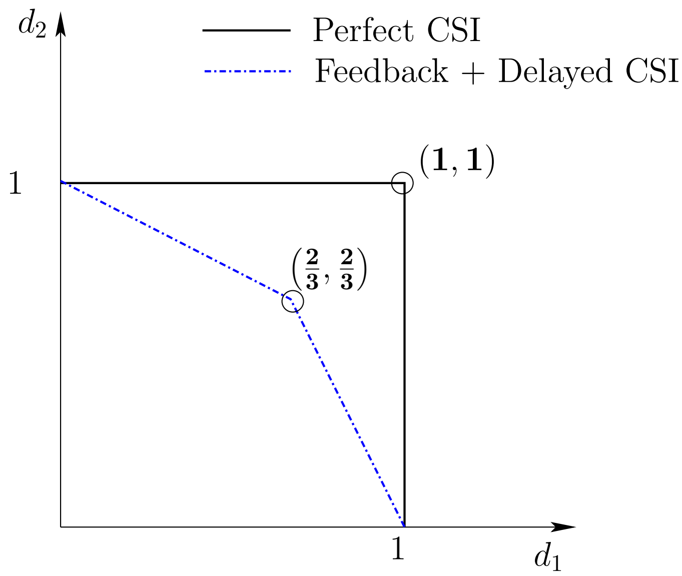

Figure 3: **DoF** regions for the $2 \times 2 \times 2$ network.

4 Coding for MIMO IC with FB and Delayed CSI

In this section, we present coding schemes that achieve the **DoF** region for MIMO IC stated in Theorem 1. We assume without loss of generality, that $N_1 \geq N_2$. We refer the reader to Table I in reference [9].

- If (M_1, M_2, N_1, N_2) are such that

$$\mathbf{DoF}^{\text{FB,d}-\text{CSI}} = \mathbf{DoF}^{\text{d}-\text{CSI}}, \tag{29}$$

  coding schemes presented in [9] which use delayed CSI only suffice for our problem. The condition (29) corresponds to cases A.I, A.II, B.0, B.I, and B.II, as defined in [9].

- If (M_1, M_2, N_1, N_2) are such that

$$\mathbf{DoF}^{\text{FB,d}-\text{CSI}} \supset \mathbf{DoF}^{\text{d}-\text{CSI}}, \tag{30}$$

  we present a novel coding scheme that achieves $\mathbf{DoF}^{\text{FB,d}-\text{CSI}}$.

We present the optimal coding schemes for the case of arbitrary (M_1, M_2, N_1, N_2)-MIMO IC in Section 7.2. In the following sub-sections, we highlight the contribution of our coding scheme through two examples which captures its essential features and leads to valuable insights for the case of general (M_1, M_2, N_1, N_2)-MIMO IC.

4.1 $(6, 2, 4, 3)$-IC with Feedback and Delayed CSI

We first focus on the case of the $(6, 2, 4, 3)$-MIMO IC. For comparison purposes, we note here the **DoF** regions with no-CSI, perfect CSI, delayed CSI, output feedback and delayed

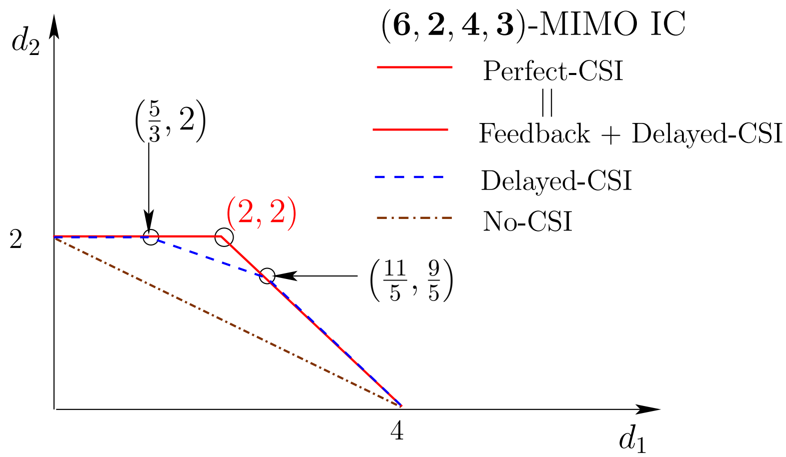

Figure 4: **DoF** region for $(6,2,4,3)$-MIMO-IC with various assumptions.

CSI. For all these four regions, we have the following bounds:

$$d_1 \leq 4; \quad d_2 \leq 2.$$

Besides these, we have the following additional bounds:

- No-CSI:

$$\frac{d_1}{4} + \frac{d_2}{2} \leq 1.$$

- Perfect CSI:

$$d_1 + d_2 \leq 4.$$

- Delayed CSI (case B-III, [9]):

$$d_1 + d_2 \leq 4; \quad d_1 + \frac{8d_2}{3} \leq 7.$$

- Output feedback and delayed CSI (Theorem 1):

$$d_1 + d_2 \leq 4; \quad \frac{d_1}{6} + \frac{d_2}{3} \leq 1.$$

It can be verified that this region is the same as the DoF region with perfect CSI, since the bound $\frac{d_1}{6} + \frac{d_2}{3} \leq 1$ is redundant as $d_1 = d_2 = 2$ is a valid choice (see Figure 4).

The main contribution of the coding scheme is to show the achievability of the point $(2,2)$ under the assumption of output feedback and delayed CSI. To show the achievability of point $(2,2)$, we will show that in three uses of the channel, we can reliably transmit 6 information symbols to receiver 1, and 6 information symbols to receiver 2.

Encoding at transmitter 2: transmitter 2 sends fresh information symbols on both its antennas for $t = 1, 2, 3$, i.e., the channel input of transmitter 2, denoted as $X_2(t)$ for $t = 1, 2, 3$ can be written as

$$X_2(1) = [v_1 \ \ v_2]^T, X_2(2) = [v_3 \ \ v_4]^T, X_2(3) = [v_5 \ \ v_6]^T.$$

At $t = 1$, transmitter 1 sends 6 information symbols on its 6 antennas, i.e., it sends

$$X_1(1) = [u_1 \ \ u_2 \ \ u_3 \ \ u_4 \ \ u_5 \ \ u_6]^T. \tag{31}$$

Let us denote by $\mathbf{u} = (u_1, \ldots, u_6)$ the vector of information symbols intended for receiver 1. The outputs at receivers 1 and 2 at $t = 1$ (ignoring noise) are given as

$$Y_1(1) = \begin{bmatrix} A_1(\mathbf{u}) + B_1(v_1, v_2) \\ A_2(\mathbf{u}) + B_2(v_1, v_2) \\ A_3(\mathbf{u}) + B_3(v_1, v_2) \\ A_4(\mathbf{u}) + B_4(v_1, v_2) \end{bmatrix}, \quad Y_2(1) = \begin{bmatrix} P_1(\mathbf{u}) + Q_1(v_1, v_2) \\ P_2(\mathbf{u}) + Q_2(v_1, v_2) \\ P_3(\mathbf{u}) + Q_3(v_1, v_2) \end{bmatrix}. \tag{32}$$

Upon receiving $Y_1(t)$ (channel output feedback) from receiver 1 and $H(1)$ (delayed CSI), transmitter 1 can use $(u_1, \ldots, u_6, Y_1(1), H(1))$ to solve for (v_1, v_2). Consequently, it can reconstruct $Q_1(v_1, v_2)$ and $Q_2(v_1, v_2)$ which constitute a part of the received signal, $Y_2(1)$, at receiver 2. In addition, having access to delayed CSI, $H(1)$, it can also compute $P_1(\mathbf{u})$ and $P_2(\mathbf{u})$, a part of the interference it caused at receiver 2. In the next two time instants, i.e., at $t = 2$ and 3, transmitter 1 sends

$$X_1(2) = [P_1(\mathbf{u}) \ \ Q_1(v_1, v_2) \ \ \phi \ \ \phi \ \ \phi \ \ \phi]^T \tag{33}$$

$$X_1(3) = [P_2(\mathbf{u}) \ \ Q_2(v_1, v_2) \ \ \phi \ \ \phi \ \ \phi \ \ \phi]^T, \tag{34}$$

where ϕ denotes a constant symbol known to all terminals.

The channel outputs at receiver 1 at $t = 2, 3$ are given as follows:

$$Y_1(2) = \begin{bmatrix} C_1(P_1(\mathbf{u}), Q_1(v_1, v_2), v_3, v_4) \\ C_2(P_1(\mathbf{u}), Q_1(v_1, v_2), v_3, v_4) \\ C_3(P_1(\mathbf{u}), Q_1(v_1, v_2), v_3, v_4) \\ C_4(P_1(\mathbf{u}), Q_1(v_1, v_2), v_3, v_4) \end{bmatrix}, \quad Y_1(3) = \begin{bmatrix} D_1(P_2(\mathbf{u}), Q_2(v_1, v_2), v_5, v_6) \\ D_2(P_2(\mathbf{u}), Q_2(v_1, v_2), v_5, v_6) \\ D_3(P_2(\mathbf{u}), Q_2(v_1, v_2), v_5, v_6) \\ D_4(P_2(\mathbf{u}), Q_2(v_1, v_2), v_5, v_6) \end{bmatrix}. \tag{35}$$

Decoding at receiver 1: Note that $Y_1(2)$ is comprised of 4 linearly independent equations in 4 variables $(P_1(\mathbf{u}), Q_1(v_1, v_2), v_3, v_4)$ and similarly, $Y_1(3)$ is comprised of 4 linearly independent equations in 4 variables $(P_2(\mathbf{u}), Q_2(v_1, v_2), v_5, v_6)$. Hence receiver 1 can decode $(P_1(\mathbf{u}), Q_1(v_1, v_2), v_3, v_4)$ from $Y_1(2)$ and it can decode $(P_2(\mathbf{u}), Q_2(v_1, v_2), v_5, v_6)$ from $Y_1(3)$. Having decoded $(Q_1(v_1, v_2), Q_2(v_1, v_2))$, it can solve for (v_1, v_2) and compute $\{B_j(v_1, v_2)\}_{j=1}^4$,

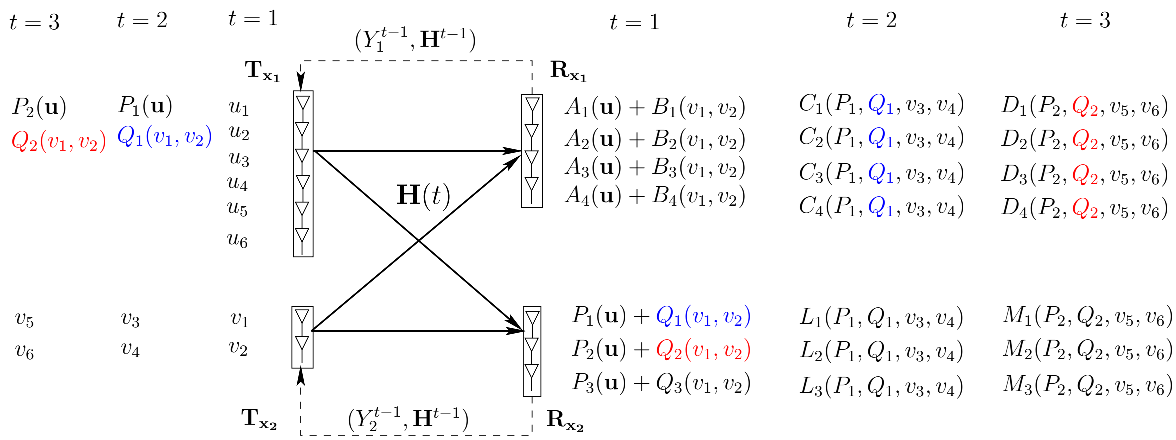

Figure 5: Coding scheme with FB, delayed CSI: $(6,2,4,3)$-MIMO-IC.

the interference caused at $t=1$. Upon subtracting $B_j(v_1,v_2)$ from the output at the jth antenna corresponding to $Y_1(1)$, receiver 1 obtains $(A_1(\mathbf{u}), A_2(\mathbf{u}), A_3(\mathbf{u}), A_4(\mathbf{u}), P_1(\mathbf{u}), P_2(\mathbf{u}))$, i.e., it has 6 linearly independent equations in 6 variables $(u_1,\ldots,u_6)$. Hence all 6 information symbols $(u_1,\ldots,u_6)$ can be decoded by receiver 1 in three uses of the channel.

The channel outputs at receiver 2 at $t=2,3$ are given as follows:

$$Y_2(2) = \begin{bmatrix} L_1(P_1(\mathbf{u}), Q_1(v_1,v_2), v_3, v_4) \\ L_2(P_1(\mathbf{u}), Q_1(v_1,v_2), v_3, v_4) \\ L_3(P_1(\mathbf{u}), Q_1(v_1,v_2), v_3, v_4) \end{bmatrix}, \quad Y_2(3) \quad = \begin{bmatrix} M_1(P_2(\mathbf{u}), Q_2(v_1,v_2), v_5, v_6) \\ M_2(P_2(\mathbf{u}), Q_2(v_1,v_2), v_5, v_6) \\ M_3(P_2(\mathbf{u}), Q_2(v_1,v_2), v_5, v_6) \end{bmatrix}. \tag{36}$$

Decoding at receiver 2: at receiver 2, we have 9 linearly independent equations (from $Y_2(1), Y_2(2), Y_2(3)$) in 9 variables $(v_1,\ldots,v_6, P_1(\mathbf{u}),\ldots,P_3(\mathbf{u}))$, where $(P_1(\mathbf{u}),\ldots,P_3(\mathbf{u}))$ is the additive interference at $t=1$. Hence it can decode 6 information symbols $(v_1,v_2,v_3,v_4,v_5,v_6)$ in three uses of the channel. Hence, we have shown the achievability of the point $(2,2)$ with channel output feedback and delayed CSI.

Remark 1: It is instructive to compare this coding scheme to the case of delayed CSI. In particular, the point $(5/3,2)$ lies on the boundary of $\mathbf{DoF}^{\mathrm{d-CSI}}$. In the coding scheme that achieves this point, it suffices to transmit 5 symbols to receiver 1 and 6 symbols to receiver 2 in three channel uses. Under the delayed CSI assumption, transmitter 1 can at best reconstruct the interference it caused at receiver 2. In the terminology of the coding scheme described above, transmitter 1 can construct $(P_1(\mathbf{u}), P_2(\mathbf{u}), P_3(\mathbf{u}))$. In subsequent channel uses, $t=2$, and $t=3$, transmitter 1 sends $P_1(\mathbf{u})$, and $P_2(\mathbf{u})$ respectively. At the end of transmission, note that receiver 2 still has 9 equations in 9 variables, $(v_1,\ldots,v_6,P_1(\mathbf{u}),\ldots,P_3(\mathbf{u}))$, therefore it can reliably decode $(v_1,\ldots,v_6)$.

The difference between the optimal coding schemes for these two models is highlighted by the decoding capability of receiver 1. For instance, in the scheme with delayed CSI alone, receiver 1 has 10 linearly independent equations in 11 variables $(u_1,\ldots,u_6,v_1,\ldots,v_6)$,

hence at best it can decode any 5 of the 6 information symbols $(u_1, \ldots, u_6)$. On the other hand, in our scheme, which allows for output feedback along with delayed CSI, transmitter 1 can exactly separate the interference and signal component of receiver 2, i.e., besides knowing $(P_1(\mathbf{u}), \ldots, P_3(\mathbf{u}))$, it can also exactly reconstruct $(Q_1(v_1, v_2), Q_2(v_1, v_2))$ (see Figure 5, which also highlights the difference of the coding schemes). This additional knowledge of $(Q_1(v_1, v_2), Q_2(v_1, v_2))$ is useful in transmission of one additional symbol to receiver 1 in three channel uses.

Remark 2: We note here that for this particular example, we can achieve the **DoF** region with perfect instantaneous CSI. Recall from [13] that the point $(2, 2)$ can be achieved with perfect CSI in one shot, i.e., in one channel use. As we have shown, output feedback and delayed CSI can also achieve the point $(2, 2)$, albeit, we pay the price of a larger delay, i.e., we can only achieve this rate in three channel uses. This observation also highlights the delay penalty incurred by the *causal* knowledge of output feedback and delayed CSI.

From this example it is clear that $\mathbf{DoF}^{\mathrm{FB,d-CSI}} = \mathbf{DoF}^{\mathrm{p-CSI}}$. However, this equality does not hold in general. In the next section, we illustrate by an example for which $\mathbf{DoF}^{\mathrm{d-CSI}} \subset \mathbf{DoF}^{\mathrm{FB,d-CSI}} \subset \mathbf{DoF}^{\mathrm{p-CSI}}$, i.e., having feedback and delayed CSI is strictly worse than having perfect CSI and strictly better than only having delayed CSI.

4.2 $(8, 4, 6, 5)$-MIMO IC

We now focus on the $(8, 4, 6, 5)$-MIMO IC (see Figure 6). The main contribution is to show the achievability of the point $(8/5, 4)$ under the assumption of output feedback and delayed CSI. To this end, we will show that in 5 channel uses, transmitter 1 can send 8 symbols to receiver 1 and transmitter 2 can send 20 symbols to receiver 2.

For all 5 channel uses, transmitter 2 sends fresh information symbols, i.e., it sends

$$X_2(1) = [v_1 \ldots v_4]^T, \ldots, X_2(5) = [v_{17} \ldots v_{20}]^T. \tag{37}$$

In the first channel use, transmitter 1 sends 8 fresh information symbols, i.e.,

$$X_1(1) = [u_1 \;\; u_2 \ldots u_8]^T. \tag{38}$$

Let us denote $\mathbf{u} = (u_1, u_2, \ldots, u_8)$. The outputs at receivers 1 and 2 at $t = 1$ (ignoring noise)

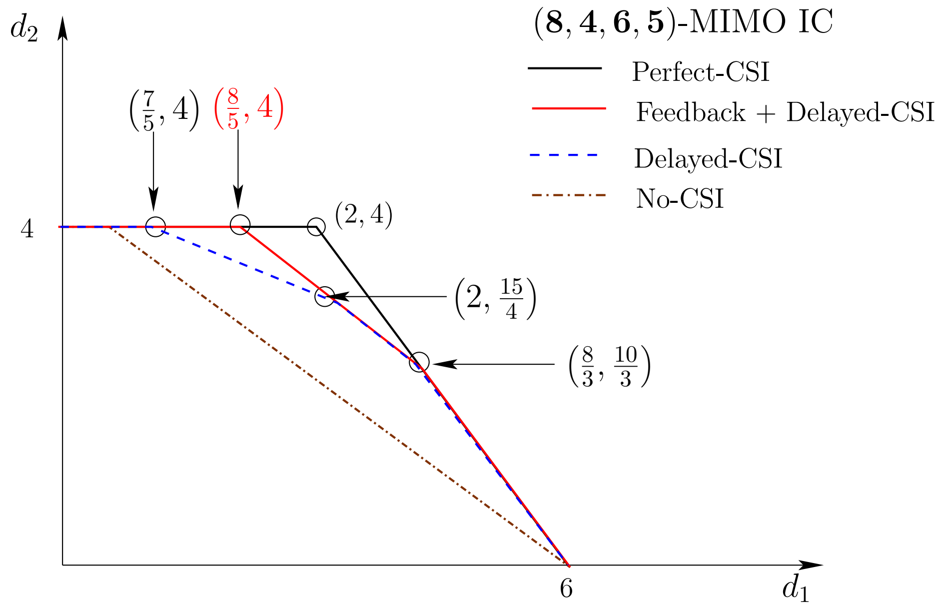

Figure 6: **DoF** region for $(8, 4, 6, 5)$-MIMO-IC with various assumptions.

are given as:

$$Y_1(1) = \begin{bmatrix} A_1(\mathbf{u}) + B_1(v_1, v_2, v_3, v_4) \\ A_2(\mathbf{u}) + B_2(v_1, v_2, v_3, v_4) \\ A_3(\mathbf{u}) + B_3(v_1, v_2, v_3, v_4) \\ A_4(\mathbf{u}) + B_4(v_1, v_2, v_3, v_4) \\ A_5(\mathbf{u}) + B_5(v_1, v_2, v_3, v_4) \\ A_6(\mathbf{u}) + B_6(v_1, v_2, v_3, v_4) \end{bmatrix}, \quad Y_2(1) = \begin{bmatrix} P_1(\mathbf{u}) + Q_1(v_1, v_2, v_3, v_4) \\ P_2(\mathbf{u}) + Q_2(v_1, v_2, v_3, v_4) \\ P_3(\mathbf{u}) + Q_3(v_1, v_2, v_3, v_4) \\ P_4(\mathbf{u}) + Q_4(v_1, v_2, v_3, v_4) \\ P_5(\mathbf{u}) + Q_5(v_1, v_2, v_3, v_4) \end{bmatrix}. \tag{39}$$

Upon receiving feedback $Y_1(1)$, and CSI $H(1)$, having access to $\mathbf{u}$, transmitter 1 can reconstruct $(P_1(\mathbf{u}), \ldots, P_4(\mathbf{u}))$ and $(Q_1(v_1, v_2, v_3, v_4), \ldots, Q_4(v_1, v_2, v_3, v_4))$. In the subsequent channel uses, $2 \leq t \leq 5$ transmitter 1 sends

$$X_1(t) = [P_{t-1}(\mathbf{u}) \;\; Q_{t-1}(v_1, v_2, v_3, v_4) \;\; \phi \;\; \phi \;\; \phi \;\; \phi \;\; \phi \;\; \phi]^T,$$

where ϕ denotes a constant symbol known to all terminals. It is straightforward to verify that receiver 2 has 25 linearly independent equations in 25 variables, $(v_1, \ldots, v_{20})$ and $(P_1(\mathbf{u}), \ldots, P_5(\mathbf{u}))$. Hence, it can decode all 20 information symbols $(v_1, \ldots, v_{20})$.

On the other hand, using $Y_1(t)$, receiver 1 can decode $P_{t-1}(\mathbf{u})$, and $Q_{t-1}(v_1, v_2, v_3, v_4)$, where $2 \leq t \leq 5$. Therefore, from $\{Y_1(t)\}_{t=2}^{5}$, it has $(P_1(\mathbf{u}), \ldots, P_4(\mathbf{u}))$ and (v_1, v_2, v_3, v_4). Using (v_1, v_2, v_3, v_4), receiver 1 can construct the interference signals $B_1(v_1, \ldots, v_4), \ldots, B_6(v_1, \ldots, v_4)$ for the first channel use. Subsequently, it can subtract these and obtain $(A_1(\mathbf{u}), \ldots, A_6(\mathbf{u}))$. To summarize, receiver 1 can obtain 10 equations $(A_1(\mathbf{u}), \ldots, A_6(\mathbf{u}), P_1(\mathbf{u}), \ldots, P_4(\mathbf{u}))$ in 8 variables and it can reliably decode $(u_1, \ldots, u_8)$.

Remark 3: From Figure 6, note that with perfect CSI, the pair $(2, 4)$ is achievable; in other words, in 5 channel uses, one can send 10 symbols to receiver 1 and 20 symbols to receiver 2. However, with output feedback and delayed CSI, to guarantee the decodability of 20 symbols at receiver 2 necessitates transmitter 1 to repeat the interference component $(P_1, \ldots, P_4)$ and a part of the signal component $(Q_1, \ldots, Q_4)$. This coding scheme fills up all the dimensions (for this example, there are 25) at receiver 2. However, this leaves 2 dimensions redundant at receiver 1, which is the reason why feedback and delayed CSI cannot achieve the point $(2, 4)$.

5 Coding for $2 \times 2 \times 2$ Network with Feedback and Delayed CSI

To achieve $4/3$ degrees of freedom, we propose a coding scheme that can transmit $4K$ symbols in $3K + 3$ uses of the channel. This scheme operates over $K + 1$ blocks, where each block is comprised of 3 channel uses, i.e., a total of $(3K + 3)$ channel uses.

We denote the symbols intended to be decoded at receiver 1 as follows:

$$(u_1^{(1)}, u_2^{(1)}), (u_1^{(2)}, u_2^{(2)}), \ldots, (u_1^{(K)}, u_2^{(K)}), \tag{40}$$

and the symbols to be decoded at receiver 2 as

$$(v_1^{(1)}, v_2^{(1)}), (v_1^{(2)}, v_2^{(2)}), \ldots, (v_1^{(K)}, v_2^{(K)}). \tag{41}$$

Here, $(u_1^{(k)}, u_2^{(k)})$ denotes the two information symbols corresponding to the k-th block intended for receiver 1, and $(v_1^{(k)}, v_2^{(k)})$ denotes the two information symbols corresponding to the k-th block intended for receiver 2.

We shall show that at the end of block k, $k > 1$, receiver 1 is able to decode the all symbols corresponding to blocks $(1, \ldots, k-1)$, i.e., it is able to decode $(u_1^{(1)}, u_2^{(1)}), \ldots, (u_1^{(k-1)}, u_2^{(k-1)})$, and similarly, receiver 2 is able to decode $(v_1^{(1)}, v_2^{(1)}), \ldots, (v_1^{(k-1)}, v_2^{(k-1)})$. Figure 7 shows a generic block k, $1 < k < K$.

5.1 Transmission in Block 1

In the first block $k = 1$; the transmitters send fresh information symbols in the first two channel uses, i.e.,

$$X_1(1) = u_1^{(1)}, \quad X_1(2) = u_2^{(1)} \tag{42}$$

$$X_2(1) = v_1^{(1)}, \quad X_2(2) = v_2^{(1)}. \tag{43}$$

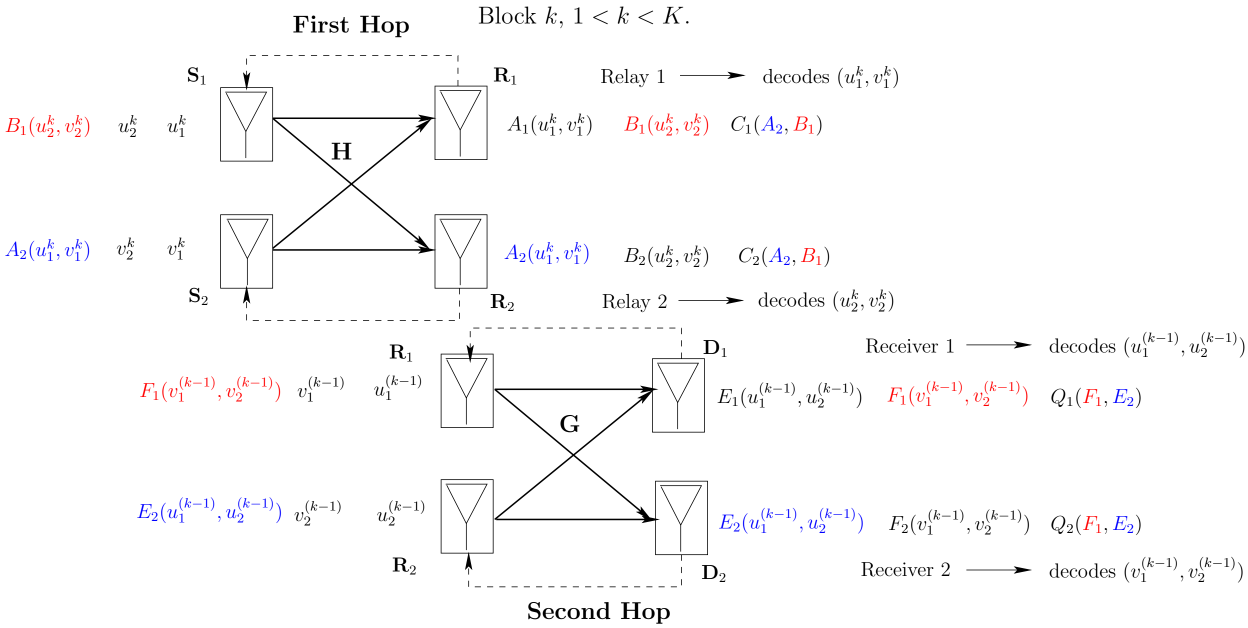

Figure 7: Coding scheme with FB and delayed CSI: $2 \times 2 \times 2$ Network

Let us consider the outputs at the relays:

$$Y_{R1}(1) = A_1(u_1^{(1)}, v_1^{(1)}), \quad Y_{R1}(2) = B_1(u_2^{(1)}, v_2^{(1)}) \tag{44}$$

$$Y_{R2}(1) = A_2(u_1^{(1)}, v_1^{(1)}), \quad Y_{R2}(2) = B_2(u_2^{(1)}, v_2^{(1)}). \tag{45}$$

Through output feedback alone, transmitter 1 knows $B_1(u_2^{(1)}, v_2^{(1)})$ and transmitter 2 knows $A_2(u_1^{(1)}, v_1^{(1)})$. In the third channel use, the transmitters send:

$$X_1(3) = B_1(u_2^{(1)}, v_2^{(1)}) \tag{46}$$

$$X_2(3) = A_2(u_1^{(1)}, v_1^{(1)}), \tag{47}$$

so that the outputs at the relays in the third channel use are

$$Y_{R1}(3) = C_1(A_2(u_1^{(1)}, v_1^{(1)}), B_1(u_2^{(1)}, v_2^{(1)})) \tag{48}$$

$$Y_{R2}(3) = C_2(A_2(u_1^{(1)}, v_1^{(1)}), B_1(u_2^{(1)}, v_2^{(1)})). \tag{49}$$

Hence, at the end of block 1, relay 1 can decode $(u_1^{(1)}, v_1^{(1)})$ and relay 2 can decode $(u_2^{(1)}, v_2^{(1)})$. In the first block, the relays remain silent.

5.2 Transmission in Block 2

The operation of the relays in this block is as follows:

$$X_{R1}(1) = u_1^{(1)}, \quad X_{R1}(2) = v_1^{(1)} \tag{50}$$

$$X_{R2}(1) = u_2^{(1)}, \quad X_{R2}(2) = v_2^{(1)}. \tag{51}$$

Let us consider the outputs at the receivers:

$$Y_1(1) = E_1(u_1^{(1)}, u_2^{(1)}), \quad Y_1(2) = F_1(v_1^{(1)}, v_2^{(1)}) \tag{52}$$

$$Y_2(1) = E_2(u_1^{(1)}, u_2^{(1)}), \quad Y_2(2) = F_2(v_1^{(1)}, v_2^{(1)}). \tag{53}$$

Through output feedback alone, relay 1 knows $F_1(v_1^{(1)}, v_2^{(1)})$ and relay 2 knows $E_2(u_1^{(1)}, u_2^{(1)})$. In the third channel use, the relays send:

$$X_{R1}(3) = F_1(v_1^{(1)}, v_2^{(1)}) \tag{54}$$

$$X_{R2}(3) = E_2(u_1^{(1)}, u_2^{(1)}), \tag{55}$$

so that the outputs at the receivers are

$$Y_1(3) = Q_1(F_1(v_1^{(1)}, v_2^{(1)}), E_2(u_1^{(1)}, u_2^{(1)})) \tag{56}$$

$$Y_2(3) = Q_2(F_1(v_1^{(1)}, v_2^{(1)}), E_2(u_1^{(1)}, u_2^{(1)})). \tag{57}$$

Hence, at the end of block 2, from $Y_1(1), Y_1(2), Y_1(3)$, receiver 1 can decode $E_1(u_1^{(1)}, u_2^{(1)})$ and $E_2(u_1^{(1)}, u_2^{(1)})$, and proceed to decode $(u_1^{(1)}, u_2^{(1)})$. Similarly, receiver 2 can decode $(v_1^{(1)}, v_2^{(1)})$.

In the second block, transmitters perform the exact same operation as in block 1, with $(u_1^{(1)}, u_2^{(1)})$ replaced by $(u_1^{(2)}, u_2^{(2)})$ and $(v_1^{(1)}, v_2^{(1)})$ replaced by $(v_1^{(2)}, v_2^{(2)})$. At the end of the second block, relay 1 can decode $(u_1^{(2)}, v_1^{(2)})$ and relay 2 can decode $(u_2^{(2)}, v_1^{(2)})$.

5.3 Transmission in Block k, $k > 2$

Just before the block k commences, relays 1 has decoded 2 symbols and relay 2 has decoded 2 symbols from the $(k-1)$th block. In general, relay 1 decodes $(u_1^{(k-1)}, v_2^{(k-1)})$ and relay 2 is able to decode $(u_2^{(k-1)}, v_1^{(k-1)})$. In block k, the relays use the exact scheme by swapping the order of transmission, i.e., relays send the u-symbols in the first channel use, v-symbols in the second channel use and the linear combinations (received via feedback) in the third channel use. In particular, at the end of block $k > 1$, the receivers are able to decode 2 symbols each, i.e., receiver 1 is able to decode $(u_1^{(k-1)}, u_2^{(k-1)})$ and receiver 2 is able to decode $(v_1^{(k-1)}, v_2^{(k-1)})$. In the very last block, the transmitters remain silent and the relays transmit. Thus, a total of $4K$ symbols can be transmitted in $(3K+3)$ channel uses, i.e., as $K \to \infty$, we can achieve a total **DoF** of 4/3. In contrast, for the case of perfect CSI, we can achieve

2 **DoF** for this model [4].

Remark 4: We note that the coding over the first hop in any given block resembles coding over an X channel with feedback [7]. In particular, each relay is able to decode 2 symbols, one from transmitter 1 and the second from transmitter 2. In the second hop, this creates a virtual $(2,1,1)$-MISO broadcast channel from the relays to the two receivers. To see this, we note that in the second hop, for block $k > 1$, the relays simultaneously send the u-symbols (intended for receiver 1) in the first channel use and send the v-symbols (intended for receiver 2) in the second channel use. In the third channel use, the relays only require feedback from their respective receivers (we note that our scheme does not require any CSI of the second hop at the relays, not even delayed CSI) to transmit linear combinations of the u symbols and v symbols.

6 Conclusions

In this paper, the **DoF** region of the MIMO-IC has been characterized under the assumption of output feedback and delayed CSI. It has been shown that output feedback and delayed CSI always outperform delayed CSI and can sometimes be as good as perfect CSI. The **DoF** region of the $2 \times 2 \times 2$ interference network has also been characterized with output feedback and delayed CSI. It has been shown that in contrast to the case of perfect CSI, for which the optimal total **DoF** is 2, in this model, the optimal total **DoF** reduces to $4/3$.

Several important and interesting open questions emerge out of this work. The usefulness of feedback in the absence of delayed CSI for both MIMO BC and the MIMO IC channel models is largely unanswered. Indeed, there are cases in which output feedback alone can achieve the same performance as feedback and delayed CSI. For instance, using output feedback alone, the achievability of $4/3$ **DoF** for the two-receiver $(2,1,1)$-MISO BC can be easily established. However, a complete characterization of the **DoF** regions for the MIMO IC and the MIMO BC channels with output feedback alone are open problems and part of our planned future work.

7 Appendix

7.1 Proof of bound (5)

The proof of this bound closely follows the technique in [9]. While proving this bound, we shall highlight the key differences for the model in consideration. We have the following

bound for the rate R_2:

$$TR_2 = H(W_2) \tag{58}$$
$$= H(W_2|\mathbf{H}^T) \tag{59}$$
$$= I(W_2; Y_2^T|\mathbf{H}^T) + H(W_2|Y_2^T, \mathbf{H}^T) \tag{60}$$
$$\leq I(W_2; Y_2^T|\mathbf{H}^T) + T\epsilon_{2,T} \tag{61}$$
$$= h(Y_2^T|\mathbf{H}^T) - h(Y_2^T|W_2, \mathbf{H}^T) \tag{62}$$
$$= Q_1 - Q_2 + T\epsilon_{2,T}. \tag{63}$$

We also have

$$TR_1 \leq I(W_1; Y_1^T, Y_2^T|W_2, \mathbf{H}^T) + T\epsilon_{1,T} \tag{64}$$
$$= h(Y_1^T, Y_2^T|W_2, \mathbf{H}^T) - h(Y_1^T, Y_2^T|W_1, W_2, \mathbf{H}^T)) + T\epsilon_{1,T} \tag{65}$$
$$= Q_3 - \sum_{t=1}^{T} h(Y_1(t), Y_2(t)|W_1, W_2, \mathbf{H}^T, Y_1^{t-1}, Y_2^{t-1}) + T\epsilon_{1,T} \tag{66}$$
$$= Q_3 - \sum_{t=1}^{T} h(Y_1(t), Y_2(t)|X_1(t), X_2(t), W_1, W_2, \mathbf{H}^T, Y_1^{t-1}, Y_2^{t-1}) + T\epsilon_{1,T} \tag{67}$$
$$= Q_3 - \sum_{t=1}^{T} h(Z_1(t), Z_2(t)) + T\epsilon_{1,T} \tag{68}$$
$$= Q_3 - To(\log_2 P) + T\epsilon_{1,T} \tag{69}$$
$$= Q_3 + T\epsilon_{1,T}, \tag{70}$$

where (67) follows from the fact that $X_1(t)$ is a function of $(W_1, \mathbf{H}^{t-1}, Y_1^{t-1})$ and $X_2(t)$ is a function of $(W_2, \mathbf{H}^{t-1}, Y_2^{t-1})$ and (68) follows from the fact that the additive noise random variables (Z_{1t}, Z_{2t}) at time t are independent of the random variables $(X_1(t), X_2(t), W_1, W_2, \mathbf{H}^T, Y_1^{t-1}, Y_2^{t-1})$.

Let $\eta(\cdot)$ denote the operation

$$\eta(x) = \lim_{P\to\infty} \left\{ \frac{1}{\log_2 P} \cdot \lim_{T\to\infty} \frac{x}{T} \right\}. \tag{71}$$

The operator $\eta(\cdot)$ is such that if $x \leq y$, then $\eta(x) \leq \eta(y)$.

Applying $\eta(\cdot)$ to both sides of (63) and (70), we obtain

$$d_2 \leq \eta(Q_1) - \eta(Q_2) \tag{72}$$
$$d_1 \leq \eta(Q_3). \tag{73}$$

From the trivial MIMO bound for a point-to-point channel, we also have

$$\eta(Q_1) \leq \min(N_2, M_1 + M_2), \tag{74}$$

which implies that

$$d_2 \leq \min(N_2, M_1 + M_2) - \eta(Q_2) \tag{75}$$

$$d_1 \leq \eta(Q_3). \tag{76}$$

The key technical step in obtaining the bound (5) is to relate the quantities $\eta(Q_2)$ and $\eta(Q_3)$ under the assumption of channel output feedback and delayed CSI. In particular, we show that

$$\eta(Q_2) \geq \frac{\min(N_2, M_1)}{\min(N_1 + N_2, M_1)} \eta(Q(3)). \tag{77}$$

As in [9], we denote

$$m_1 = \min(N_1 + N_2, M_1) \tag{78}$$

$$m_2 = \min(N_2, M_1) \tag{79}$$

Hence, proving (77) is equivalent to proving

$$\eta(Q_2) \geq \frac{m_2}{m_1} \eta(Q(3)). \tag{80}$$

To this end, we first focus on the term Q_2:

$$\begin{aligned} Q_2 &= h(Y_2^T | \mathbf{H}^T, W_2) && (81) \\ &= \sum_{t=1}^{T} h(Y_2(t) | W_2, \mathbf{H}^T, Y_2^{t-1}) && (82) \\ &= \sum_{t=1}^{T} h(Y_2(t) | X_2(t), W_2, \mathbf{H}^T, Y_2^{t-1}), && (83) \end{aligned}$$

where (83) follows from the fact that $X_2(t)$ is a function of $(W_2, \mathbf{H}^{t-1}, Y_2^{t-1})$.

Notation: Through the rest of this section, we denote $Y_1^{[1:a]}(t)$ as the vector of outputs received on the upper most a antennas at receiver 1 at time t. Furthermore, $Y_{1j}(t)$ denotes the (scalar) output received on the jth antenna at receiver 1 at time t.

We now focus on the t-th term in the summation in (83),

$$\begin{aligned}
h(Y_2(t)|X_2(t), W_2, \mathbf{H}^T, Y_2^{t-1}) &= h(Y_2^{[1:m_2]}(t), Y_2^{[m_2+1:N_2]}(t)|X_2(t), W_2, \mathbf{H}^T, Y_2^{t-1}) && (84)\\
&= h(Y_2^{[1:m_2]}(t)|X_2(t), W_2, \mathbf{H}^T, Y_2^{t-1}) \\
&\quad + h(Y_2^{[m_2+1:N_2]}(t)|X_2(t), W_2, \mathbf{H}^T, Y_2^{t-1}, Y_2^{[1:m_2]}(t)) && (85)\\
&\geq h(Y_2^{[1:m_2]}(t)|X_2(t), W_2, \mathbf{H}^T, Y_2^{t-1}) \\
&\quad + h(Y_2^{[m_2+1:N_2]}(t)|X_1(t), X_2(t), W_2, \mathbf{H}^T, Y_2^{t-1}, Y_2^{[1:m_2]}(t)) && (86)\\
&= h(Y_2^{[1:m_2]}(t)|X_2(t), W_2, \mathbf{H}^T, Y_2^{t-1}) + h(Z_2^{[m_2+1:N_2]}(t)) && (87)\\
&= h(Y_2^{[1:m_2]}(t)|X_2(t), W_2, \mathbf{H}^T, Y_2^{t-1}) + o(\log_2(P)) && (88)\\
&= h(Y_2^{[1:m_2]}(t)|X_2(t), \mathbf{H}(t), W_2, \mathbf{H}^T \setminus \mathbf{H}(t), Y_2^{t-1}) + o(\log_2(P)) && (89)\\
&= h(Y_2^{[1:m_2]}(t)|X_2(t), \mathbf{H}(t), U(t)) + o(\log_2(P)), && (90)
\end{aligned}$$

where in (90), we have defined $U(t)$ as follows:

$$U(t) \triangleq (W_2, \mathbf{H}^T \setminus \mathbf{H}(t), Y_2^{t-1}). \qquad (91)$$

Hence (83) together with (90) imply that

$$Q_2 \geq \sum_{t=1}^{T} h(Y_2^{[1:m_2]}(t)|X_2(t), \mathbf{H}(t), U(t)) + T \cdot o(\log_2(P)). \qquad (92)$$

We now focus on the term Q_3:

$$\begin{aligned}
Q_3 &= h(Y_1^T, Y_2^T|W_2, \mathbf{H}^T) && (93)\\
&= \sum_{t=1}^{T} h(Y_1(t), Y_2(t)|W_2, \mathbf{H}^T, Y_1^{t-1}, Y_2^{t-1}) && (94)\\
&= \sum_{t=1}^{T} h(Y_1(t), Y_2(t)|X_2(t), W_2, \mathbf{H}^T, Y_1^{t-1}, Y_2^{t-1}) && (95)\\
&\leq \sum_{t=1}^{T} h(Y_1(t), Y_2(t)|X_2(t), W_2, \mathbf{H}^T, Y_2^{t-1}) && (96)\\
&= \sum_{t=1}^{T} h(Y_1(t), Y_2(t)|X_2(t), \mathbf{H}(t), U(t)). && (97)
\end{aligned}$$

We next have the following Lemma:

Lemma 1

$$h(Y_1(t), Y_2(t)|X_2(t), \mathbf{H}(t), U(t)) \leq h(Y_1^{[1:m_1-m_2]}(t), Y_2^{[1:m_2]}(t)|X_2(t), \mathbf{H}(t), U(t)) + o(\log_2 P). \tag{98}$$

Proof: This lemma is trivial if $M_1 \geq N_1 + N_2$. Henceforth, we assume that $M_1 < N_1 + N_2$. For this case, consider the expansion

$$\begin{aligned} & h(Y_1(t), Y_2(t)|X_2(t), \mathbf{H}(t), U(t)) \\ & = h(Y_1^{[1:m_1-m_2]}(t), Y_2^{[1:m_2]}(t)|X_2(t), \mathbf{H}(t), U(t)) \\ & \quad + h(Y_1^{[m_1-m_2+1:N_1]}(t), Y_2^{[m_2+1:N_2]}(t)|X_2(t), \mathbf{H}(t), U(t), Y_1^{[1:m_1-m_2]}(t), Y_2^{[1:m_2]}(t)). \end{aligned} \tag{99}$$

Next, note that given $(X_2(t), \mathbf{H}(t))$ and $(Y_1^{[1:m_1-m_2]}(t), Y_2^{[1:m_2]}(t))$, the contribution due to $X_2(t)$ can be subtracted off and one is left with $m_1 = \min(M_1, N_1+N_2) = M_1$ linear equations that can be solved to obtain $X_1(t)$ with probability 1. Therefore, the differential entropy of the second term in (99) equals the differential entropy of the additive noise random variables, which is constant w.r.t. P. Hence, this argument implies that

$$h(Y_1(t), Y_2(t)|X_2(t), \mathbf{H}(t), U(t)) \leq h(Y_1^{[1:m_1-m_2]}(t), Y_2^{[1:m_2]}(t)|X_2(t), \mathbf{H}(t), U(t)) + o(\log_2 P). \tag{100}$$

■

Lemma 1 together with (97) implies that

$$Q_3 \leq \sum_{t=1}^{T} h(Y_1^{[1:m_1-m_2]}(t), Y_2^{[1:m_2]}(t)|X_2(t), \mathbf{H}(t), U(t)) + T \cdot o(\log_2 P). \tag{101}$$

Recall that from (92), we have

$$Q_2 \geq \sum_{t=1}^{T} h(Y_2^{[1:m_2]}(t)|X_2(t), \mathbf{H}(t), U(t)) + T \cdot o(\log_2(P)). \tag{102}$$

We shall now prove (80) from (101) and (102). In particular, we will show that

$$Q_2 \geq \frac{m_2}{m_1} Q_3. \tag{103}$$

Expanding the bound on Q_3 in (101), we have

$$Q_3 \leq \sum_{t=1}^{T} h(Y_1^{[1:m_1-m_2]}(t), Y_2^{[1:m_2]}(t)|X_2(t), \mathbf{H}(t), U(t)) + T \cdot o(\log_2 P) \tag{104}$$

$$= \sum_{t=1}^{T} h(Y_2^{[1:m_2]}(t)|X_2(t), \mathbf{H}(t), U(t))$$

$$+ \sum_{t=1}^{T} h(Y_1^{[1:m_1-m_2]}(t)|X_2(t), \mathbf{H}(t), U(t), Y_2^{[1:m_2]}(t)) + T \cdot o(\log_2 P) \tag{105}$$

$$\leq Q_2 + \sum_{t=1}^{T} h(Y_1^{[1:m_1-m_2]}(t)|X_2(t), \mathbf{H}(t), U(t), Y_2^{[1:m_2]}(t)) \tag{106}$$

$$= Q_2 + \sum_{t=1}^{T} \left[\sum_{i=1}^{(m_1-m_2)} h(Y_{1i}(t)|X_2(t), \mathbf{H}(t), U(t), Y_2^{[1:m_2]}(t), Y_1^{[1:i-1]}(t)) \right] \tag{107}$$

$$\leq Q_2 + \sum_{t=1}^{T} \left[\sum_{i=1}^{(m_1-m_2)} h(Y_{1i}(t)|X_2(t), \mathbf{H}(t), U(t), Y_2^{[1:m_2-1]}(t)) \right] \tag{108}$$

$$= Q_2 + \sum_{t=1}^{T} \left[(m_1 - m_2) h(Y_{11}(t)|X_2(t), \mathbf{H}(t), U(t), Y_2^{[1:m_2-1]}(t)) \right] \tag{109}$$

where (106) follows from (102), and (108) follows from the fact that conditioning reduces differential entropy. The last step (109) follows from the statistical equivalence of the signals at any two distinct antennas. We prove this claim as follows. Consider any $i \neq j$,

$$h(Y_{1i}(t)|X_2(t), \mathbf{H}(t), U(t), Y_2^{[1:m_2-1]}(t))$$

$$= \mathbb{E}_{H_{11i}(t)=\mathbf{a}} \left[h(Y_{1i}(t)|X_2(t), \mathbf{H}(t-1), U(t), Y_2^{[1:m_2-1]}(t), H_{11i}(t) = \mathbf{a}) \right] \tag{110}$$

$$= \mathbb{E}_{H_{11j}(t)=\mathbf{a}} \left[h(Y_{1j}(t)|X_2(t), \mathbf{H}(t-1), U(t), Y_2^{[1:m_2-1]}(t), H_{11j}(t) = \mathbf{a}) \right] \tag{111}$$

$$= h(Y_{1j}(t)|X_2(t), \mathbf{H}(t), U(t), Y_2^{[1:m_2-1]}(t)), \tag{112}$$

where (111) follows from the fact that given $(X_2(t), \mathbf{H}(t-1), U(t), Y_2^{[1:m_2-1]}(t))$, the joint distributions of the random variables $(H_{11i}(t), X_1(t))$ and $(H_{11j}(t), X_1(t))$ are identical.

Now, consider the bound on Q_2 in (102):

$$Q_2 \geq \sum_{t=1}^{T} h(Y_2^{[1:m_2]}(t)|X_2(t), \mathbf{H}(t), U(t)) + T \cdot o(\log_2(P)) \tag{113}$$

$$= \sum_{t=1}^{T} \left[\sum_{i=1}^{m_2} h(Y_{2i}(t)|X_2(t), \mathbf{H}(t), U(t), Y_2^{[1:i-1]}(t)) \right] + T \cdot o(\log_2(P)) \tag{114}$$

We next have the following claim for $1 \leq i \leq (m_2 - 1)$:

$$\begin{aligned}
& h(Y_{2i}(t)|X_2(t), \mathbf{H}(t), U(t), Y_2^{[1:i-1]}(t)) \\
& = h(Y_{2(i+1)}(t)|X_2(t), \mathbf{H}(t), U(t), Y_2^{[1:i-1]}(t)) && (115)\\
& \geq h(Y_{2(i+1)}(t)|X_2(t), \mathbf{H}(t), U(t), Y_2^{[1:i]}(t)), && (116)
\end{aligned}$$

where (115) follows by same argument as we used to prove (111), and (116) follows from the fact that conditioning reduces differential entropy. Using this sequence of inequalities, we can lower bound (114) as follows:

$$\begin{aligned}
Q_2 &\geq \sum_{t=1}^{T} h(Y_2^{[1:m_2]}(t)|X_2(t), \mathbf{H}(t), U(t)) + T \cdot o(\log_2(P)) && (117)\\
&= \sum_{t=1}^{T} \left[\sum_{i=1}^{m_2} h(Y_{2i}(t)|X_2(t), \mathbf{H}(t), U(t), Y_2^{[1:i-1]}(t)) \right] + T \cdot o(\log_2(P)) && (118)\\
&\geq \sum_{t=1}^{T} \left[m_2 h(Y_{2m_2}(t)|X_2(t), \mathbf{H}(t), U(t), Y_2^{[1:m_2-1]}(t)) \right] + T \cdot o(\log_2(P)) && (119)\\
&= \sum_{t=1}^{T} \left[m_2 h(Y_{11}(t)|X_2(t), \mathbf{H}(t), U(t), Y_2^{[1:m_2-1]}(t)) \right] + T \cdot o(\log_2(P)), && (120)
\end{aligned}$$

where (119) follows from the sequence of inequalities in (116), and (120) follows from the statistical equivalence of the (scalar) outputs at any two distinct antennas.

Collecting the bounds in (109) and (120), we have shown that

$$Q_3 \leq Q_2 + (m_1 - m_2) \sum_{t=1}^{T} \left[h(Y_{11}(t)|X_2(t), \mathbf{H}(t), U(t), Y_2^{[1:m_2-1]}(t)) \right], \tag{121}$$

and

$$Q_2 \geq m_2 \sum_{t=1}^{T} \left[h(Y_{11}(t)|X_2(t), \mathbf{H}(t), U(t), Y_2^{[1:m_2-1]}(t)) \right] + T \cdot o(\log_2(P)). \tag{122}$$

From these two bounds, we have

$$Q_3 \leq Q_2 + \frac{(m_1 - m_2)}{m_2} Q_2 + \frac{T \cdot o(\log_2(P))}{m_2} \tag{123}$$

which is equivalent to

$$Q_3 \leq \frac{m_1}{m_2} Q_2 + \frac{T \cdot o(\log_2(P))}{m_2}. \tag{124}$$

By taking the operator $\eta(\cdot)$ on both sides, we have the proof that

$$\eta(Q_2) \geq \frac{m_2}{m_1}\eta(Q_3). \tag{125}$$

Now the proof of (5) is straightforward from

$$d_2 \leq \min(N_2, M_1 + M_2) - \eta(Q_2) \tag{126}$$

$$d_1 \leq \eta(Q_3), \tag{127}$$

which together with (125) imply

$$\begin{aligned}
\min(N_2, M_1 + M_2) &\geq d_2 + \eta(Q_2) && (128)\\
&\geq d_2 + \frac{m_2}{m_1}\eta(Q_3) && (129)\\
&\geq d_2 + \frac{m_2}{m_1}d_1 && (130)\\
&= d_2 + \frac{\min(N_2, M_1)}{\min(N_1 + N_2, M_1)}d_1, && (131)
\end{aligned}$$

which is equivalent to the desired bound (5):

$$\frac{d_1}{\min(N_1 + N_2, M_1)} + \frac{d_2}{\min(N_2, M_1)} \leq \frac{\min(N_2, M_1 + M_2)}{\min(N_2, M_1)}. \tag{132}$$

7.2 Coding Scheme: arbitrary (M_1, M_2, N_1, N_2)

We focus on only such values of (M_1, M_2, N_1, N_2) for which $\mathbf{DoF}^{\mathrm{d-CSI}} \subset \mathbf{DoF}^{\mathrm{FB,d-CSI}}$. A necessary condition for this inclusion is $M_1 > N_1 > N_2 > M_2$. For this case, the region in Theorem 1 can be simplified to

$$d_2 \leq M_2 \tag{133}$$

$$d_1 + d_2 \leq N_1 \tag{134}$$

$$\frac{d_1}{\min(M_1, N_1 + N_2)} + \frac{d_2}{N_2} \leq 1. \tag{135}$$

We can further subdivide this scenario into two mutually exclusive cases, depending on whether the bound (135) is active or not:

7.2.1 $\mathbf{DoF}^{\mathrm{FB,d-CSI}} = \mathbf{DoF}^{\mathrm{p-CSI}}$

In this case, the bound (135) is not active and hence the region is the same as that of perfect CSI. This condition requires (M_1, M_2, N_1, N_2) to satisfy

$$\frac{(N_1 - M_2)}{\min(M_1, N_1 + N_2)} + \frac{M_2}{N_2} \leq 1, \tag{136}$$

which is equivalent to

$$\min(M_1, N_1 + N_2) \geq N_2 \left(\frac{N_1 - M_2}{N_2 - M_2} \right) \tag{137}$$

The **DoF** region with feedback and delayed CSI is given as:

$$d_2 \leq M_2 \tag{138}$$

$$d_1 + d_2 \leq N_1 \tag{139}$$

The main contribution is to show the achievability of the following point:

$$\text{Point } P_0 : (N_1 - M_2, M_2). \tag{140}$$

The $(6, 2, 4, 3)$-MIMO IC falls in this category.

7.2.2 $\mathbf{DoF}^{\mathrm{FB,d-CSI}} \subset \mathbf{DoF}^{\mathrm{p-CSI}}$

In this case, the bound (135) is active and hence the region with feedback and delayed CSI is a strict subset of the **DoF** region with perfect CSI. This condition requires (M_1, M_2, N_1, N_2) to satisfy

$$\frac{(N_1 - M_2)}{\min(M_1, N_1 + N_2)} + \frac{M_2}{N_2} > 1, \tag{141}$$

which is equivalent to

$$\min(M_1, N_1 + N_2) > N_2 \left(\frac{N_1 - M_2}{N_2 - M_2} \right) \tag{142}$$

The $(8, 4, 6, 5)$-MIMO IC falls in this category.

The **DoF** region with feedback and delayed CSI is given as:

$$d_2 \leq M_2 \tag{143}$$

$$d_1 + d_2 \leq N_1 \tag{144}$$

$$\frac{d_1}{\min(M_1, N_1 + N_2)} + \frac{d_2}{N_2} \leq 1. \tag{145}$$

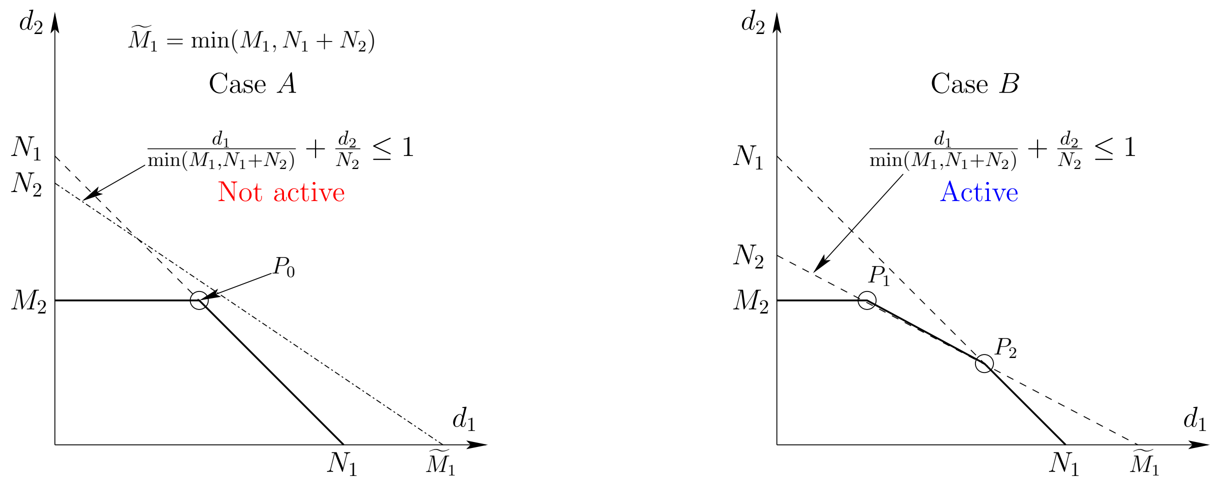

Figure 8: Two cases for (M_1, M_2, N_1, N_2)-MIMO-IC.

The main contribution is to show the achievability of two points:

$$\text{Point } P_1 : \left[\widetilde{M}_1 \left(\frac{N_2 - M_2}{N_2}\right), M_2\right], \tag{146}$$

$$\text{Point } P_2 : \left[\widetilde{M}_1 \left(\frac{N_1 - N_2}{\widetilde{M}_1 - N_2}\right), N_2 \left(\frac{\widetilde{M}_1 - N_1}{\widetilde{M}_1 - N_2}\right)\right], \tag{147}$$

where we have defined $\widetilde{M}_1 = \min(M_1, N_1 + N_2)$.

We have shown these two cases in figure 8.

Achievability for P_0 and P_1: Let us define

$$L = \max\left(\frac{\widetilde{M}_1}{N_1 - M_2}, \frac{N_2}{N_2 - M_2}\right). \tag{148}$$

We will provide a scheme that works well for both P_0 (corresponding to Case A) and for Point P_1 (corresponding to Case B). In particular, we will show that in L channel uses, we can transmit $\widetilde{M}_1$ symbols to receiver 1 and LM_2 symbols to receiver 2. Note that the technical condition distinguishing cases A and B can be equivalently stated in terms of the value taken by the parameter L.

In this scheme, transmitter 2 always sends fresh information symbols in all L channel uses. Transmitter 1 at $t = 1$ sends $\widetilde{M}_1$ fresh information symbols. After $t = 1$, receiver 2 has N_2 equations in M_2 information symbols and N_2 interference components (which are functions of $\tilde{M}_1$ symbols of transmitter 1). Via feedback from receiver 1, and having delayed CSI, transmitter 1 can exactly recover these M_2 information symbols and M_2 interference components it caused at receiver 2. In the subsequent $(L - 1)$ channel uses, transmitter 1 sends combinations of these symbols on its antennas.

At receiver 2, we have LN_2 equations in $LM_2 + N_2$ variables. Hence, for the decodability

of LM_2 symbols at receiver 2, L must satisfy $LN_2 \geq LM_2 + N_2$, i.e.,

$$L \geq \frac{N_2}{N_2 - M_2}. \tag{149}$$

Furthermore, at receiver 1, we have LN_1 equations in $\widetilde{M}_1 + LM_2$ variables, and hence for the decodability of $\widetilde{M}_1$ symbols at receiver 1, L must satisfy $LN_1 \geq \widetilde{M}_1 + LM_2$, i.e.,

$$L \geq \frac{\widetilde{M}_1}{N_1 - M_2}. \tag{150}$$

Note that we choose this exact value of L in (148) to ensure the decoding requirements at both the decoders. Consequently, we have shown the achievability of points P_0 and P_1.

Achievability for P_2: Let us define

$$L = \widetilde{M}_1 - N_2. \tag{151}$$

We will show that in L channel uses, we can transmit $\widetilde{M}_1(N_1 - N_2)$ symbols to receiver 1 and $N_2(\widetilde{M}_1 - N_1)$ symbols to receiver 2. Before proceeding, we verify the feasibility of such a scheme. Note that transmitter 2 has LM_2 total number of antennas (over L channel uses) to send fresh information. Hence, for such a scheme to work, this number should exceed the total number of information symbols to be sent to receiver 2, i.e., we must have

$$LM_2 \geq N_2(\widetilde{M}_1 - N_1), \tag{152}$$

which is equivalent to

$$\widetilde{M}_1 \geq N_2 \left(\frac{N_1 - M_1}{N_2 - M_1} \right). \tag{153}$$

This condition is clearly satisfied from (142) and the fact that $M_1 > M_2$.

We now propose the coding scheme for point P_2: transmitter 2 sends $N_2(\widetilde{M}_1 - N_1)$ information symbols in L channel uses. Transmitter 1 sends fresh information on $\widetilde{M}_1$ antennas in the first $(N_1 - N_2)$ channel uses (note that $L > (N_1 - N_2)$). From the first $(N_1 - N_2)$ channel uses, upon receiving feedback and delayed CSI, transmitter 1 can reconstruct the $M_2(N_1 - N_2)$ information components and $N_2(N_1 - N_2)$ interference components of receiver 2. In the subsequent $L - (N_1 - N_2)$ channel uses, transmitter 1 forwards these two components using M_1 antennas.

Decoding at receiver 2: at the end of transmission, receiver 2 has access to LN_2 linearly independent equations in $N_2(\widetilde{M}_1 - N_1)$ information symbols and $N_2(N_1 - N_2)$ interference

components. Thus, for receiver 2 to decode the information symbols, we must have

$$LN_2 \geq N_2(\widetilde{M}_1 - N_1) + N_2(N_1 - N_2) \tag{154}$$

$$= N_2(\widetilde{M}_1 - N_2). \tag{155}$$

Decoding at receiver 1: receiver 1 has LN_1 equations in $\widetilde{M}_1(N_1 - N_2)$ information symbols and $N_2(\widetilde{M}_1 - N_1)$ interference symbols. Therefore, for decoding at receiver 1 to succeed, we must have

$$LN_1 \geq \widetilde{M}_1(N_1 - N_2) + N_2(\widetilde{M}_1 - N_1) \tag{156}$$

$$= N_1(\widetilde{M}_1 - N_2). \tag{157}$$

Indeed we have chosen L in (151) to satisfy both (155) and (157) with equality. Thus, we have shown the achievability of the point P_2 with feedback and delayed CSI.

References

[1] M.A. Maddah-Ali, A.S. Motahari, and A.K. Khandani. Communication over MIMO X channels: Interference alignment, decomposition, and performance analysis. *IEEE Transactions on Information Theory*, 54(8):3457–3470, 2008.

[2] V.R. Cadambe and S.A. Jafar. Interference alignment and degrees of freedom of the K-user interference channel. *IEEE Transactions on Information Theory*, 54(8):3425–3441, 2008.

[3] S. Mohajer, S. Diggavi, C. Fragouli, and D. Tse. Approximate capacity characterization for a class of Gaussian relay-interference wireless networks. *IEEE Transactions on Information Theory*, 57(5):2837–2864, May 2011.

[4] T. Gou, S.A. Jafar, S.W. Jeon, and S.Y. Chung. Aligned interference neutralization and the degrees of freedom of the 2x2x2 interference channel. *Arxiv preprint arXiv:1012.2350*, 2010.

[5] Syed A. Jafar. Interference alignment: A new look at signal dimensions in a communication network. *Foundations and Trends in Communications and Information Theory*, 7(1):1–134, 2010.

[6] M.A. Maddah-Ali and D. Tse. Completely stale transmitter channel state information is still very useful. In *48th Annual Allerton Conference on Communication, Control, and Computing, 2010*, 2010.

[7] H. Maleki, S.A. Jafar, and S. Shamai. Retrospective interference alignment. *Arxiv preprint arXiv:1009.3593*, 2010.

[8] H. Maleki, S.A. Jafar, and S. Shamai. Retrospective interference alignment. In *Proc. IEEE International Symposium on Information Theory (ISIT)*, St. Petersburg, Russia, 2011.

[9] C.S. Vaze and M.K. Varanasi. The degrees of freedom region and interference alignment for the MIMO interference channel with delayed CSI. *IEEE Trans. on Information Theory,* submitted, January 2011.

[10] C.S. Vaze and M.K. Varanasi. The degrees of freedom region of the MIMO interference channel with delayed CSIT. In *Proc. IEEE International Symposium on Information Theory (ISIT)*, St. Petersburg, Russia, 2011.

[11] C. Suh and D. Tse. Feedback capacity of the Gaussian interference channel to within two bits. *IEEE Trans. on Information Theory*, 57(5):2667–2685, May 2011.

[12] C.S. Vaze and M.K. Varanasi. The degrees of freedom region of the two-user MIMO broadcast channel with delayed CSI. *Arxiv preprint arXiv:1101.0306*, 2010.

[13] C. Huang and S.A. Jafar. Degrees of freedom of the MIMO interference channel with cooperation and cognition. *IEEE Trans. on Information Theory*, 55(9):4211–4220, Sep. 2009.